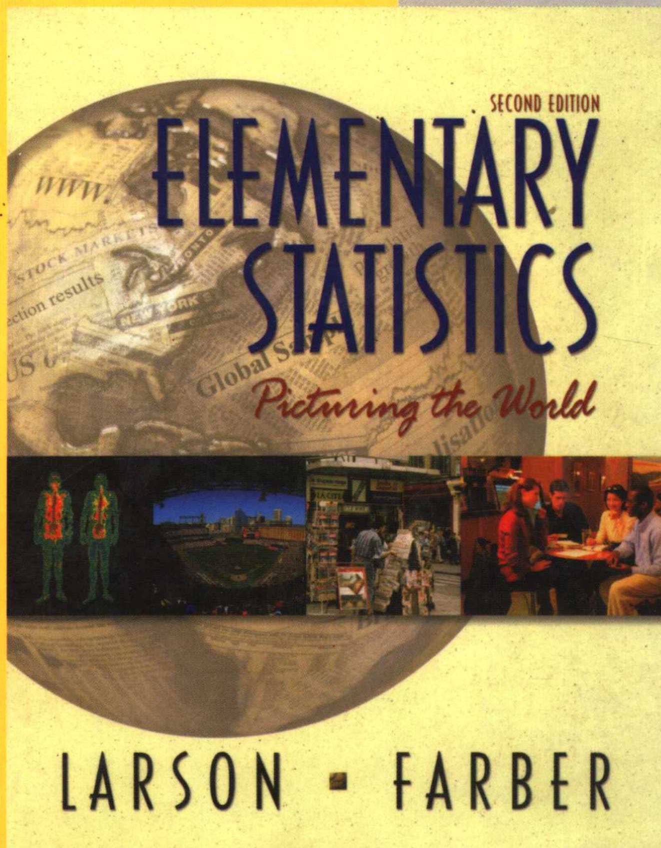
PEARSON
Prentice Hall

国外大学优秀教材——统计学系列（影印版）

Ron Larson, Betsy Farber

基础统计学

（第2版）



Pearson Education

清华大学出版社

国外大学优秀教材——统计学系列（影印版）

基础统计学

（第2版）

Elementary Statistics
—Picturing the World
(Second Edition)

Ron Larson
Betsy Farber

清华大学出版社
北京

EISBN: 0-13-065595-3

图书在版编目(CIP)数据

基础统计学=Elementary Statistics—Picturing the World: 第 2 版/拉森(Larson, R.),法伯(Farber, B.)著.—影印本.—北京:清华大学出版社,2004.11
(国外大学优秀教材.统计学系列)
ISBN 7-302-09723-2

Ⅰ.基… Ⅱ.①拉…②法… Ⅲ.统计学—高等学校—教材—英文 Ⅳ.C8

中国版本图书馆 CIP 数据核字(2004)第 106003 号

出 版 者: 清华大学出版社　　**地　址**: 北京清华大学学研大厦
http://www.tup.com.cn　　**邮　编**: 100084
社 总 机: 010-62770175　　**客户服务**: 010-62776969

责任编辑: 刘晓艳
印 刷 者: 北京市清华园胶印厂
装 订 者: 北京国马印刷厂
发 行 者: 新华书店总店北京发行所
开　本: 210×248　**印张**: 44.75
版　次: 2004 年 11 月第 1 版　2004 年 11 月第1 次印刷
书　号: ISBN 7-302-09723-2/O · 416
印　数: 1～3000
定　价: 69.00 元

国外大学优秀教材——统计学系列
（影印版）

出版说明

从科学研究、工农业生产、新产品开发、产品质量的提高到政治、教育、社会科学等各个领域,使用统计方法和不使用统计方法获得的结果是大不相同的。只要统计方法使用得当,就能够得到事半功倍的效果。这也是统计学能在经济发达国家兴旺发达的原因。

在我国的高等教育中,统计学的教学内容和国外先进水平相比还有一定的差距,统计学的研究和应用还没有得到足够的重视,统计学的方法还不为更多的应用工作者所了解。为了尽快地改进统计学的教学内容,促进统计学在我国的研究和应用,我们将陆续引进一批在国外知名大学受到普遍欢迎的统计学教材。希望对我国统计学的高等教育、科学研究和应用做出一点贡献。

这批教材的引进将遵循以下的基本原则:适合作为我国大学本科或研究生统计学课程的教材或主要教学参考书;原则上是近年国际上出版的最新图书或新版图书;对于某些基础课程,选择不同体系、不同风格和不同层次的教材,以满足不同层次和不同学时的需要。每本入选教材都有国内相应领域知名专家或资深教授的审阅和推荐。对于部分教学需要量较大的教材还将考虑用翻译的方式引进。统计学系列教材将分期分批出版,内容将覆盖统计学专业大多数基础课和选修课。

概率论是和统计学密切相关的学科,也是学习研究统计学的基础,所以概率论方面的教材也是这次引进的重点。我们将注重引进一批应用性强,和统计学联系紧密的优秀概率论教材,以满足各方面的需要。

本系列教材的内容将包含概率统计、随机过程、Bayes 统计、多元统计、数据分析等,同时兼顾金融经济统计、社会统计、商业统计、生物统计、质量控制、实验设计、市场调查等等。系列教材的读者是统计学专业及需要学习统计学课程和知识的其他各专业的本科生、研究生以及从事与概率统计学科相关工作的科研工作者和应用工作者。

我们希望本系列教材的引进能够对我国统计学的高等教育、科学研究和实际应用有所帮助。也希望得到广大读者的反馈意见,以便改进我们的工作。

何书元
北京大学数学学院

Elementary Statistics—Picturing the World(Second Edition)

影印版序

由 Ron Larson 和 Betsy Farber 撰写的教材“Elementary Statistics—Picturing the World”出版于2000 年，并于 2003 出版了第 2 版。作者广泛地使用案例分析的方法，通过大量的应用实例，引导人们如何正确地收集数据资料，进行统计分析，以得到有意义的参考性结论。

本书的内容包括数据的收集、整理与描述，概率的概念，基本的统计方法等。各章编排如下：第 1 章统计引言。简单扼要地通过案例介绍了什么是统计，不同类型的数据，数据的采集及有关的试验设计。第 2 章描述性统计方法。这是一种利用图形和图表对数据进行粗加工的简单易行的方法。它可以在很短的时间内，使人们对于数据所提供的信息有一个综观的了解。这种方法在各种展览资料和广告资料中被频繁地使用。第 3 章概率的基本概念。第 4 章离散概率分布。第 5 章正态概率分布。它完全通过图形引入，避免给出解析表达式，从而避免使用微积分的知识。有关的计算都是由统计软件完成的。第 6 章置信区间。第 7 章单样本假设检验。第 8 章双样本假设检验。第 9 章相关与回归。第 10 章 χ^2 检验与 F 分布，包括列联表、单因素和双因素方差分析。第 11 章非参数检验。

本书是典型的文科教材，适合作为语言文学、工商管理、人文科学、社会科学、心理科学、考古学以及经济类有关专业的大学生或研究生的统计学教材。

本书在撰写上强调了以下特点：

(1) 起点低，使用面广，实用性强，强调统计思想的内涵与应用。读者只需具有高中数学知识就可以学习使用本书。在本书中，不用微积分，不追求数学公式的推导与形式逻辑思维的推理，而代之以在应用中不断地使用这些公式与运用形象思维和直观判断。

(2) 在编排上，本书以案例为中心展开，全书贯穿以对数据的统计分析，充分利用图表分析阐述统计思想，从中引出统计结论。书中针对每一个案例都总结了关键概念，便于读者复习提高。

(3) 本书是自封闭的，在书中包含了概率及分布的精要内容，因而不需要读者先学概率论的知识。

(4) 各章安排了许多学习小技巧和注解、对比应用与滥用的例子，以及简单统计软件的使用方法等内容。在各章末有小结及统计实例与决策实例，并配置了大量在各个社会科学领域中的应用习题，还配有许多测试题。本书所附的光盘中提供了与各章有关的统计分析计算的原始数据，这些数据分别以以下三种格式存储：普通文本格式、Microsoft Excel 格式、统计软件 Minitab 格式。由于 Microsoft Excel 在台式电脑与手提电脑中是普遍装载的，这对于使用者特别方便；而 Minitab 具有更多的功能，适合于扩展应用，它是一个相对小而相当全的实用统计软件。

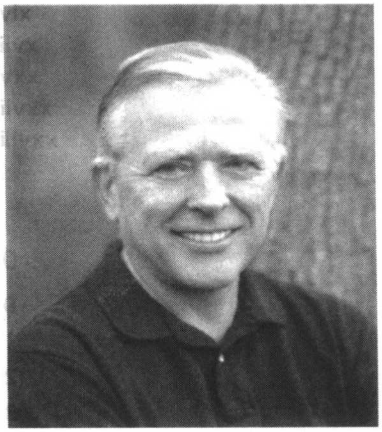
(5) 本书在内容上概括了直观统计学和常用统计方法的基本内容。即使对已有大学数学水平的教师和学生，通过浏览此书，也能在直觉推断和形象思维方面有所裨益。此书也为进一步学习较为深入的统计思想与方法提供了台阶。

当前，在我国的人文科学与社会科学等广大领域中，科学地用统计分析作决策的要求与日俱增，此书则提供了学习使用统计思维和作实际统计推断的各种方法。

本书语言流畅生动，很适合于双语教学，也适合决策人员、财贸人员、农林畜牧等有关领域的教师、工程师、技术人员参阅或自学，成为他们提高科学分析与管理能力的实用工具与良师益友。

龚光鲁
清华大学数学科学系

About the Authors



Ron Larson

Penn State University at Erie

Ron Larson received his Ph.D. in mathematics from the University of Colorado in 1970. At that time he accepted a position with Penn State University, and he currently holds the rank of professor of mathematics at the university. Larson is the lead author of more than two dozen mathematics textbooks that range from sixth grade through calculus levels. Many of his texts, such as the seventh edition of his calculus text, are leaders in their markets. Larson is also one of the pioneers in the use of multimedia and the Internet to enhance the learning of mathematics. He has authored multimedia programs, extending from the elementary school through calculus levels. Larson is a member of several professional groups and is a frequent speaker at national and regional mathematics meetings.

Betsy Farber

Bucks County Community College

Betsy Farber received her Bachelor's degree in mathematics from Penn State University and Master's degree in mathematics from the College of New Jersey. Since 1976, she has been teaching all levels of mathematics at Bucks County Community College in Newtown, Pennsylvania, where she currently holds the rank of professor. She is particularly interested in developing new ways to make statistics relevant and interesting to her students, and has been teaching statistics in many different modes—with *TI-83*, with *MINITAB*, and by distance learning as well as in the traditional classroom. A member of the American Mathematical Association of Two-Year Colleges (AMATYC), she is an author of *The Student Edition to MINITAB* and *A Guide to MINITAB.* She served as consulting editor for *Statistics, A First Course* and has written computer tutorials for the CD-ROM correlating to the texts in the Streeter Series in mathematics.

Preface

Welcome to *Elementary Statistics: Picturing the World*, Second Edition. We are grateful for the overwhelming acceptance and support of the First Edition. It is gratifying to know that our vision of combining theory, pedagogy, and design to exemplify how statistics is used to picture and describe the world has helped students learn about statistics and make informed decisions. This message—picturing the world—begins with the cover and is carefully integrated into every feature of the text.

New to the Second Edition

Two features in the Second Edition help students apply statistics to real-life situations and practice making decision about statistics.

Uses and Abuses Each chapter now has a full page summarizing the uses of concepts in the chapter, as well as a description of common misuses. Each "abuse" is accompanied by one or more exercises.

Real Statistics–Real Decisions—Putting It All Together Following the **Review Exercises** in each chapter, we have added a full-page real-life situation accompanied by exercises that ask students to use the concepts in the chapter to make decisions.

The exercise sets in the Second Edition include approximately 200 new exercises, giving the students more practice in performing calculations, making decisions, providing explanations, and applying results to a real-life setting.

In response to suggestions from statistics instructors, the coverage of topics in Chapters 2, 5, 7, and 9 is revised in the Second Edition.

- In **Chapter 2,** the z-score is now introduced in Section 2.5, **Measures of Position.**
- In **Chapter 5,** we added two new sections—Section 5.3, **Normal Distributions: Finding Probabilities,** and Section 5.4, **Normal Distributions: Finding Values.** These sections replace Section 5.3, **Applications of Normal Distributions** in the First Edition. Changing these sections allows the instructor to cover applications of normal distributions in greater detail and from two perspectives.
- In **Chapter 7,** Section 7.1, **Introduction to Hypothesis Testing,** we now introduce the concept of hypothesis testing using probability values, or P-values. The concept of using rejection regions is now introduced in Section 7.2, **Hypothesis Testing for the Mean (Large Samples).**
- In **Chapter 9,** for instructors who prefer to cover Section 9.1, **Correlation,** immediately after covering graphing paired data in Chapter 2, we added a method for testing a population correlation coefficient that does not involve hypothesis testing. The method is simple and can easily be covered after Chapters 1 and 2.

General Features

Versatile Course Coverage The table of contents of the text was developed to give instructors **many options.** For instance, by assigning the **Extending the Basics** exercises and spending time on the chapter projects, there is sufficient content to use the text in a two-semester course. More commonly, we expect the text to be used in a three-credit semester course or a four-semester course that includes a lab component. In such cases, instructors will have to pare down the text's 46 sections. If you want more information on sample syllabi, check the Web site that accompanies the text, *www.prenhall.com/larson*.

Choice of Tables Our experience has shown that students find a **cumulative density function** (CDF) table easier to use than a "0-to-z" table. Using the CDF table to find the area under a normal curve is the topic of Section 5.2 on pages 214–218. Because we realize that many teachers prefer to use the "0-to-z" table, we have provided an alternative presentation of Section 5.2 using the "0-to-z" table in Appendix A of the book.

Graphical Approach As with most introductory statistics texts, we begin the descriptive statistics chapter with a survey of different ways to display data graphically. A difference between this text and many others is that **we continue to incorporate the graphical display of data throughout the text.** For example, see the use of stem-and-leaf plots to display data on pages 348 and 349. In all, the text has over 900 graphs—surpassing all other introductory statistics texts.

Variety of Real-Life Applications We have chosen real-life applications that are representative of the majors of the students taking introductory statistics courses. These include business, psychology, health sciences, sports, computer science, political science, and many others. Choosing meaningful applications for such a diverse audience is difficult. We wanted the applications to be **authentic**—but they also need to be **accessible.**

Data and Source Lines The data sets in the book were chosen for interest, variety, and their ability to illustrate concepts. Most of the **over 200 data sets** contain actual data with source lines. The remaining data sets contain simulated data that, though not actual, are representative of real-life situations. All data sets containing 20 or more entries are available in a variety of electronic forms, including disk and Internet. In the exercise sets, the data sets that are available electronically are indicated by the icon .

Accuracy Every effort was made to **ensure the mathematical accuracy** of the examples and exercise solutions. The examples and exercises were solved by two people independently. A third person compared the independent solutions and resolved differences. If you encounter errors that we missed, please contact us so that we can correct the problem in a subsequent printing.

Balanced Approach The text strikes a **balance between computation, decision making, and conceptual understanding.** We have provided many Examples, Exercises, and Try It problems that go beyond mere computation. For instance, look at Exercises 31 and 32 on page 45. Students are not just asked to construct a relative frequency histogram for the given data, they are asked to go a step further and use the histogram to make a decision.

Prerequisites Statistics contains many formulas and variables, including radicals, summation notation, Greek letters, and subscripts. So, **some familiarity with algebra and evaluation of algebraic expressions** is a prerequisite. Nevertheless, we have made every effort to keep algebraic manipulations to a minimum—often we display informal versions of formulas using words in place of or in addition to variables. For instance, see the definitions of midpoint and relative frequency on page 34.

Flexible Technology Although most formulas in the book are illustrated with tabular "hand" calculations, we assume that most students who take this course have access to some form of technology tool, such as *MINITAB*, *Excel*, the *TI-83*, or *SPSS*. Because the use of technology varies widely, we have made the text flexible. **It can be used in courses with no more technology than a scientific calculator—or it can be used in courses that require frequent use of sophisti-cated technology tools.** For those who want specific instructions on particular technology tools, separate technology manuals are available to augment the text. Whatever your use of technology, we are sure that you agree with us that the goal of this course is not computation. Rather, it is to gain an understanding of the basic concepts and uses of statistics.

Page Layout We believe that statistics is more accessible to students when it is carefully formatted on each page with a consistent open layout. This text is the **first college-level statistics book to be written to design,** which means that its features (Examples, Try It problems, Definitions, or Guidelines) are not split from one page to the next. Although this process requires extra planning and work in the development stage, the result is a presentation that is clean and clear.

MAA, AMATYC, NCTM Standards This text answers the call for a **student-friendly text that emphasizes the uses of statistics** and not just the computation of its myriad of formulas. Our experience indicates that our job as instructors of an introductory course in statistics is not to produce statisticians but to produce informed consumers of statistical reports. For this reason, we have included many exercises that require students to provide written explanations, find patterns, and make decisions.

Features

Chapter Openers

Where You've Been

Each chapter begins with a two-page photographic description of a real-life problem. The first page has a feature called *Where You've Been*. It shows students how the chapter fits into the bigger picture of statistics, by connecting it to topics learned in earlier chapters.

Where You've Been

In Chapters 1 through 5, you studied descriptive statistics (how to collect and describe data) and probability (how to find probabilities and analyze discrete and continuous probability distributions). For instance, the Wheat Quality Council uses descriptive statistics to analyze the data collected during its annual crop tour.

In a recent crop tour, 444 wheat fields were sampled. Of the 288 fields of spring wheat, the mean yield was 32.5 bushels per acre with a standard deviation of 10.9 bushels per acre. Of the 156 fields of durum wheat, the mean yield was 26.8 bushels per acre with a standard deviation of 8.0 bushels per acre.

Donn Pikop, a buyer for a milling company, checks a wheat spike in a Minnesota field.

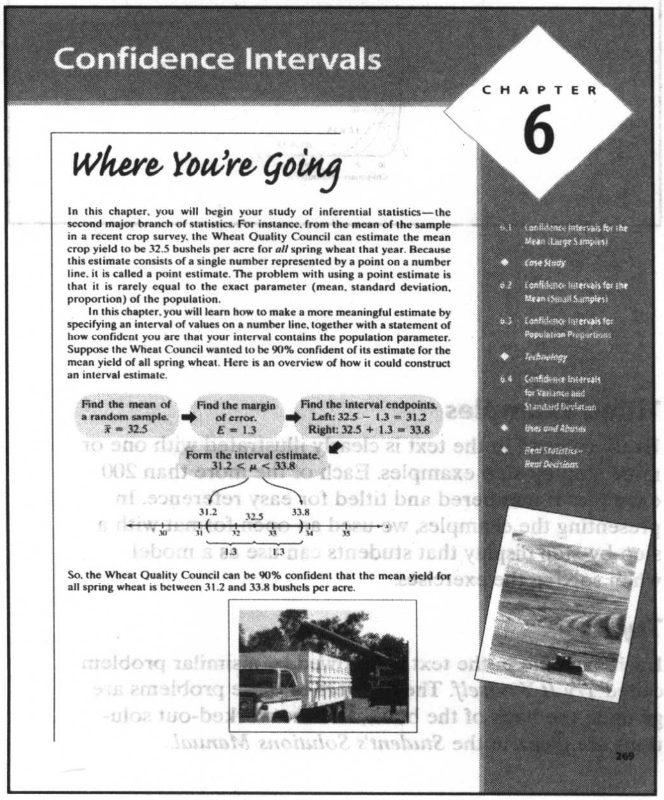
The Wheat Quality Council has its headquarters in Pierre, South Dakota. Its primary function is to encourage the development and production of new and better varieties of wheat. The Council also assesses the yield and quality of wheat crops. For instance, in a recent crop survey, the wheat averaged about 61 pounds per bushel with a 14% protein content.

268

Chapter Openers

Where You're Going

The second page of the chapter opener has a feature called *Where You're Going*. It gives students an overview of the chapter, exploring concepts in the context of real-world settings.



Confidence Intervals

CHAPTER 6

Where You're Going

In this chapter, you will begin your study of inferential statistics—the second major branch of statistics. For instance, from the mean of the sample in a recent crop survey, the Wheat Quality Council can estimate the mean crop yield to be 32.5 bushels per acre for *all* spring wheat that year. Because this estimate consists of a single number represented by a point on a number line, it is called a point estimate. The problem with using a point estimate is that it is rarely equal to the exact parameter (mean, standard deviation, proportion) of the population.

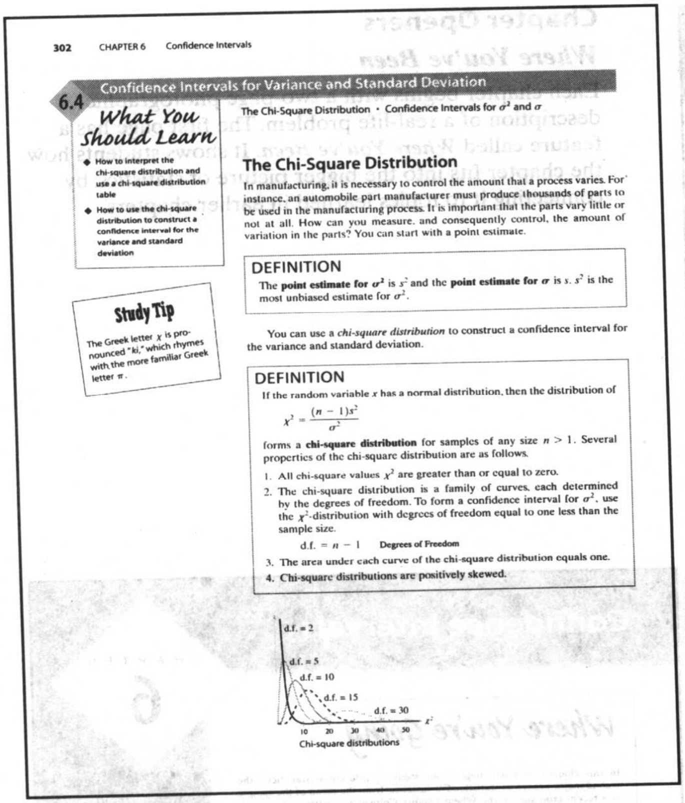
In this chapter, you will learn how to make a more meaningful estimate by specifying an interval of values on a number line, together with a statement of how confident you are that your interval contains the population parameter. Suppose the Wheat Council wanted to be 90% confident of its estimate for the mean yield of all spring wheat. Here is an overview of how it could construct an interval estimate.

Find the mean of a random sample. $\bar{x} = 32.5$ → Find the margin of error. $E = 1.3$ → Find the interval endpoints. Left: $32.5 - 1.3 = 31.2$ Right: $32.5 + 1.3 = 33.8$ → Form the interval estimate. $31.2 < \mu < 33.8$

So, the Wheat Quality Council can be 90% confident that the mean yield for all spring wheat is between 31.2 and 33.8 bushels per acre.

6.1 Confidence Intervals for the Mean (Large Samples)
◆ Case Study
6.2 Confidence Intervals for the Mean (Small Samples)
6.3 Confidence Intervals for Population Proportions
◆ Technology
6.4 Confidence Intervals for Variance and Standard Deviation
◆ Uses and Abuses
◆ Real Statistics—Real Decisions

269



302 CHAPTER 6 Confidence Intervals

6.4 Confidence Intervals for Variance and Standard Deviation

What You Should Learn

- How to interpret the chi-square distribution and use a chi-square distribution table
- How to use the chi-square distribution to construct a confidence interval for the variance and standard deviation

The Chi-Square Distribution · Confidence Intervals for σ^2 and σ

The Chi-Square Distribution

In manufacturing, it is necessary to control the amount that a process varies. For instance, an automobile part manufacturer must produce thousands of parts to be used in the manufacturing process. It is important that the parts vary little or not at all. How can you measure, and consequently control, the amount of variation in the parts? You can start with a point estimate.

DEFINITION

The **point estimate for σ^2** is s^2 and the **point estimate for σ** is s. s^2 is the most unbiased estimate for σ^2.

Study Tip

The Greek letter χ is pronounced "ki," which rhymes with the more familiar Greek letter π.

You can use a *chi-square distribution* to construct a confidence interval for the variance and standard deviation.

DEFINITION

If the random variable x has a normal distribution, then the distribution of

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

forms a **chi-square distribution** for samples of any size $n > 1$. Several properties of the chi-square distribution are as follows.

1. All chi-square values χ^2 are greater than or equal to zero.
2. The chi-square distribution is a family of curves, each determined by the degrees of freedom. To form a confidence interval for σ^2, use the χ^2-distribution with degrees of freedom equal to one less than the sample size.
   d.f. $= n - 1$ **Degrees of Freedom**
3. The area under each curve of the chi-square distribution equals one.
4. **Chi-square distributions are positively skewed.**

Section Organization

Each section is organized by learning objectives. These objectives are presented in everyday language in a margin feature called *What You Should Learn*. The same objectives are then used as subsection titles throughout the section.

Study Tips

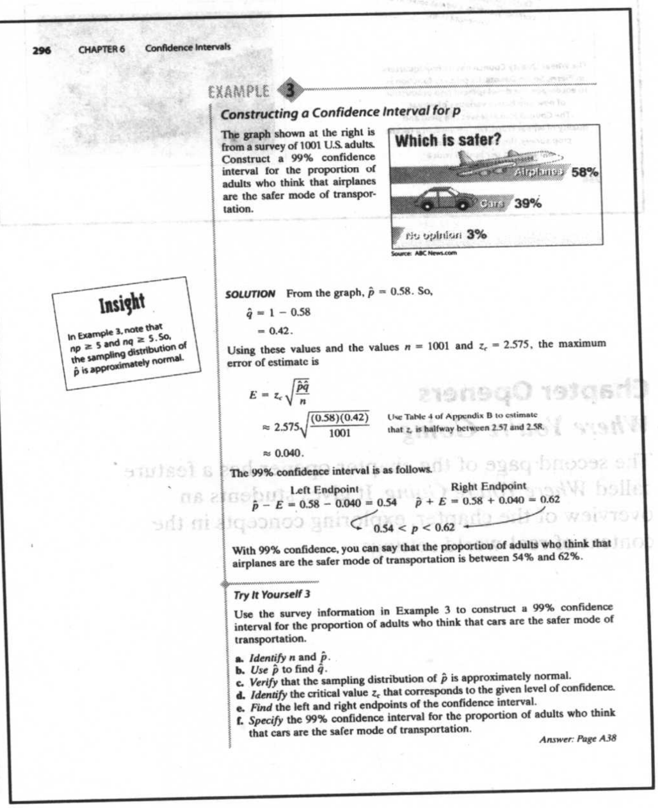
Most sections contain one or more study tips placed on yellow "sticky notes" in the margin. These tend to be informal learning aids, which show how to read a table, use technology, or interpret a result or a graph.

Titled Examples

Every concept in the text is clearly illustrated with one or more step-by-step examples. Each of the more than 200 examples is numbered and titled for easy reference. In presenting the examples, we used an open format with a step-by-step display that students can use as a model when solving the exercises.

Try Its

Each example in the text is followed by a similar problem called *Try It Yourself*. The answers to these problems are given in the back of the book, and the worked-out solutions are given in the *Student's Solutions Manual*.



296 CHAPTER 6 Confidence Intervals

EXAMPLE 3

Constructing a Confidence Interval for p

The graph shown at the right is from a survey of 1001 U.S. adults. Construct a 99% confidence interval for the proportion of adults who think that airplanes are the safer mode of transportation.

Insight

In Example 3, note that $np \geq 5$ and $nq \geq 5$. So, the sampling distribution of $\hat{p}$ is approximately normal.

SOLUTION From the graph, $\hat{p} = 0.58$. So,

$$\hat{q} = 1 - 0.58 = 0.42.$$

Using these values and the values $n = 1001$ and $z_c = 2.575$, the maximum error of estimate is

$$E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}} \approx 2.575\sqrt{\frac{(0.58)(0.42)}{1001}} \approx 0.040.$$

Use Table 4 of Appendix B to estimate that z_c is halfway between 2.57 and 2.58.

The 99% confidence interval is as follows.

Left Endpoint: $\hat{p} - E = 0.58 - 0.040 = 0.54$ Right Endpoint: $\hat{p} + E = 0.58 + 0.040 = 0.62$

$$0.54 < p < 0.62$$

With 99% confidence, you can say that the proportion of adults who think that airplanes are the safer mode of transportation is between 54% and 62%.

Try It Yourself 3

Use the survey information in Example 3 to construct a 99% confidence interval for the proportion of adults who think that cars are the safer mode of transportation.

a. *Identify* n and $\hat{p}$.
b. *Use* $\hat{p}$ to find $\hat{q}$.
c. *Verify* that the sampling distribution of $\hat{p}$ is approximately normal.
d. *Identify* the critical value z_c that corresponds to the given level of confidence.
e. *Find* the left and right endpoints of the confidence interval.
f. *Specify* the 99% confidence interval for the proportion of adults who think that cars are the safer mode of transportation.

Answer: Page A38

Confidence Intervals for the Population Mean

Using a point estimate and a maximum error of estimate, you can construct an interval estimate of a population parameter such as μ. This interval estimate is called a confidence interval.

DEFINITION

A c-confidence interval for the population mean μ is

$$\bar{x} - E < \mu < \bar{x} + E.$$

The probability that the confidence interval contains μ is c.

GUIDELINES

Finding a Confidence Interval for a Population Mean ($n \geq 30$ or σ known with a normally distributed population)

In Words	*In Symbols*
1. Find the sample statistics n and $\bar{x}$.	$\bar{x} = \frac{\Sigma x}{n}$
2. Specify σ, if known. Otherwise, if $n \geq 30$, find the sample standard deviation s and use it as an estimate for σ.	$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}$
3. Find the critical value z_c that corresponds to the given level of confidence.	Use the Standard Normal Table.
4. Find the maximum error of estimate E.	$E = z_c \frac{\sigma}{\sqrt{n}}$
5. Find the left and right endpoints and form the confidence interval.	Left endpoint: $\bar{x} - E$; Right endpoint: $\bar{x} + E$; Interval: $\bar{x} - E < \mu < \bar{x} + E$

Study Tip

The left and right endpoints of a confidence interval are $\bar{x} - E$ and $\bar{x} + E$, respectively. Other ways to represent a confidence interval are $(\bar{x} - E, \bar{x} + E)$ and $\bar{x} \pm E$. For instance, in Example 3, you could write the confidence interval as (11.1, 13.7) or 12.4 ± 1.3.

EXAMPLE 3

Constructing a Confidence Interval

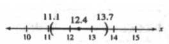
See *Minitab* steps on page 308.

Construct a 95% confidence interval for the mean number of sentences in all magazine advertisements.

SOLUTION In Examples 1 and 2, you found that $\bar{x} = 12.4$ and $E = 1.3$. The confidence interval is as follows.

Left Endpoint	Right Endpoint
$\bar{x} - E = 12.4 - 1.3 = 11.1$	$\bar{x} + E = 12.4 + 1.3 = 13.7$

$$11.1 < \mu < 13.7$$



So, with 95% confidence, you can say that the population mean number of sentences is between 11.1 and 13.7.

Definitions

The critical statistics definitions are set off with gold screens. Formal definitions are often followed by guidelines that explain, in everyday English, how to apply the definition.

Guidelines

Throughout the book, the presentation of a statistical formula is followed by a set of step-by-step guidelines for applying the formula. The guidelines are divided into two columns titled *In Words* and *In Symbols*.

Picturing the World

Two footballs, one filled with air and the other filled with helium, were kicked on a windless day at Ohio State University. The footballs were alternated with each kick. After 10 practice kicks, each football was kicked 29 more times. The distances (in yards) are listed. *(Source: OSC Scientists Get a Kick Out of Sports Controversy, "The Columbus Dispatch," November 21, 1993.)*

Air Filled

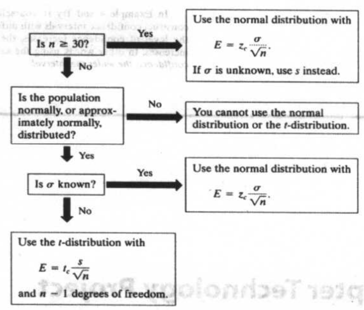
1	9
2	0 0 2 2 2
2	5 5 5 5 6 6
2	7 7 7 8 8 8 8 8 9 9 9
3	1 1 1 2
3	3 4 Key: 1|9 = 19

Helium Filled

1	1 2
1	4
1	
2	2
2	3 4 6 6 6
2	7 8 8 8 9 9 9 9
3	0 0 0 0 1 1 2 2
3	3 4 5
3	9

Assume that the distances are normally distributed for each football. Apply the flowchart at the right to each sample. Find a 95% confidence interval for the mean distance each football traveled. Do the confidence intervals overlap? What does this tell you?

The flowchart describes when to use the normal distribution to construct a confidence interval and when to use a t-distribution.



EXAMPLE 4

Choosing the Normal or t-Distribution

You randomly select 25 newly constructed houses. The sample mean construction cost is $181,000 and the population standard deviation is $28,000. Assuming construction costs are normally distributed, should you use the normal distribution, the t-distribution, or neither to construct a 95% confidence interval for the population mean construction cost? Explain your reasoning.

SOLUTION Because the population is normally distributed and the population standard deviation is known, you should use the normal distribution.

Try It Yourself 4

You randomly select 18 adult male athletes and measure the resting heart rate of each. The sample mean heart rate is 64 beats per minute with a sample standard deviation of 2.5 beats per minute. Assuming the heart rates are normally distributed, should you use the normal distribution, the t-distribution, or neither to construct a 90% confidence interval for the mean heart rate? Explain your reasoning.

a. Use the flowchart to determine which distribution you should use to construct the 90% confidence interval for the mean heart rate.

Answer: Page A38

Picturing the World

Each section contains a real-life "mini case study" that illustrates the important concept or concepts of the section. Each *Picturing the World* concludes with a question.

Technology Examples

Many sections contain a worked example that shows how technology can be used to calculate formulas, perform tests, or display data. Screen displays from *MINITAB*, *Excel*, and *TI-83* are given. Additional screen displays are given at the ends of selected chapters, and detailed instructions are given in separate technology manuals available with the book.

Insights

Most sections also contain one or more insights placed on blue "sticky notes" in the margin. The purpose of each insight is to help drive home an important interpretation or help connect different concepts.

274 CHAPTER 6 Confidence Intervals

Insight

The width of a confidence interval is $2E$. Examine the formula for E to see why a larger sample size tends to give you a narrower confidence interval for the same level of confidence.

Try It Yourself 3

Use the data given in Try It Yourself 1 to construct a 95% confidence interval for the mean number of sentences in all magazine advertisements. Compare your result to the interval found in Example 3.

a. *Find* $\bar{x}$ and E.
b. *Find* the left and right endpoints of the confidence interval.
c. *State* the 95% confidence interval for the mean number of sentences in all magazine advertisements and compare it to Example 3.

Answer: Page A37

EXAMPLE 4

Constructing a Confidence Interval Using Technology

Use a computer or graphing calculator to construct a 99% confidence interval for the mean number of sentences in all magazine advertisements, using the sample in Example 1.

SOLUTION To use a technology tool to solve the problem, enter the data and find that the sample standard deviation is $s \approx 5.0$. Then, use the confidence interval command to calculate the confidence interval (ZInterval for *TI-83*, 1-Sample Z for *Minitab*). The displays should look like the ones shown here.

TI-83

```
ZInterval
 (10.673, 14.179)
 x̄=12.42592593
 Sx=5.015454801
 n=54
```

MINITAB

Z Confidence Intervals

The assumed sigma = 5

Variable	N	Mean	StDev	SE Mean	99.0 % CI
C1	54	12.426	5.015	0.680	(10.673, 14.179)

So, a 99% confidence interval for μ is (10.7, 14.2). With 99% confidence, you can say that the population mean number of sentences is between 10.7 and 14.2.

Try It Yourself 4

Use the sample data in Example 1 and a computer or graphing calculator to construct 75% and 85% confidence intervals for the mean number of sentences in all magazine advertisements. How does the width of the confidence interval change as the level of confidence increases?

a. *Enter* the data.
b. *Use* the appropriate command to construct each confidence interval.
c. *Compare* the widths of the confidence intervals for $c = 0.75, 0.85$, and 0.99.

Answer: Page A37

In Example 4 and Try It Yourself 4, the same sample data was used to construct confidence intervals with different levels of confidence. Notice that as the level of confidence increases, the width of the confidence interval also increases. In other words, using the same sample data, ***the greater the level of confidence, the wider the interval.***

Chapter Technology Project

Each chapter has a full-page technology projects using tools from *MINITAB*, *Excel*, and *TI-83*, that gives students additional insight into the way technology is used to handle large data sets or complex, real-life questions.

Technology 301

TECHNOLOGY | MINITAB | EXCEL | TI-83

WWW.GALLUP.COM

THE GALLUP ORGANIZATION

Most Admired Polls

From 1948 to 2000 the Gallup Organization has conducted a "most admired" poll. The methodology for the 2000 poll is described at the right.

Survey Question

What man that you have heard or read about, living today in any part of the world, do you admire most? And who is your second choice?*

*Survey respondents are asked an identical question about most admired woman.

"The results are based on telephone interviews with a randomly selected national sample of 1011 adults, 18 and over, conducted December 15–17, 2000. For results based on samples of this size, one can say with 95 percent confidence that the error attributable to sampling and other random effects could be plus or minus 3 percentage points. In addition to sampling error, question wording and practical difficulties in conducting surveys can introduce error or bias into the findings of public opinion polls."

Exercises

1. In 2000, 61 people named Pope John Paul II as their most admired man. Use a technology tool to find a 95% confidence interval for the proportion that would have chosen the pope.
2. Does the confidence interval you obtained in Exercise 1 agree with the statement issued by the Gallup Organization that the proportion is 6% plus or minus 3%? Explain.
3. The most named woman was Hillary Clinton. The second named woman was Oprah Winfrey, who was named by 4% of the people in the sample. Use a technology tool to find a 95% confidence interval for the proportion of the population that would have chosen Oprah Winfrey.
4. Use a technology tool to simulate a most admired poll. Assume that the actual population proportion who most admire Oprah Winfrey is 7%. Run the simulation several times using $n = 1011$.

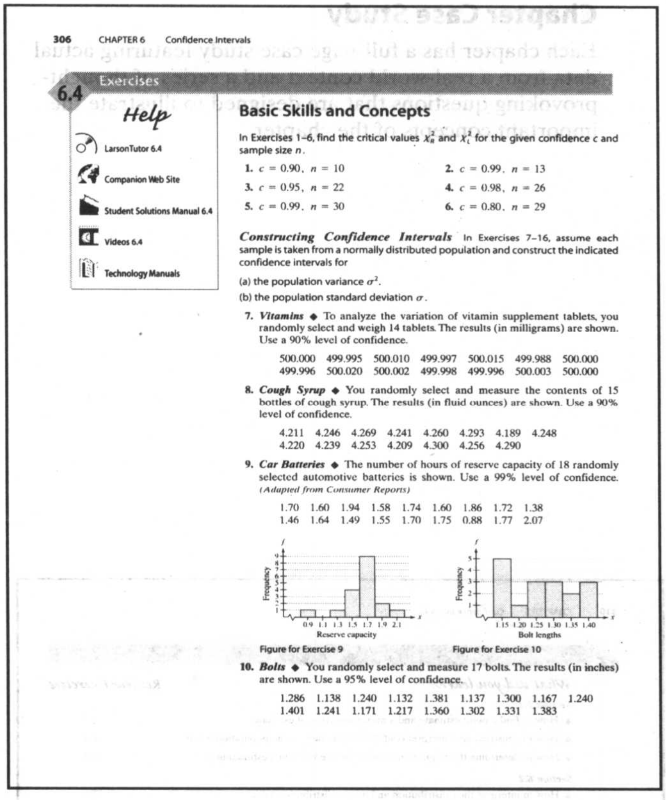
(a) What was the least value you obtained for $\hat{p}$?
(b) What was the greatest value you obtained for $\hat{p}$?

MINITAB

```
Generate 200 rows of data
Store in column(s):  C1
Number of trials: 1011
Probability of success:  .07
```

5. The 2000 results were a rarity because in the poll's 52-year history, it was unusual that no one (other than the first lady) received more than 6% of the vote. Is it possible, however, that the actual proportion of the population that most admired Oprah Winfrey was 7% or greater? Explain your reasoning.

Extended solutions are given in the *Technology Supplement*.
Technical instruction is provided for *Minitab*, *Excel*, and the *TI-83*.



306 CHAPTER 6 Confidence Intervals

6.4 Exercises

Help

LarsonTutor 6.4

Companion Web Site

Student Solutions Manual 6.4

Videos 6.4

Technology Manuals

Basic Skills and Concepts

In Exercises 1–6, find the critical values χ_R^2 and χ_L^2 for the given confidence c and sample size n.

1. $c = 0.90,\ n = 10$
2. $c = 0.99,\ n = 13$
3. $c = 0.95,\ n = 22$
4. $c = 0.98,\ n = 26$
5. $c = 0.99,\ n = 30$
6. $c = 0.80,\ n = 29$

Constructing Confidence Intervals In Exercises 7–16, assume each sample is taken from a normally distributed population and construct the indicated confidence intervals for

(a) the population variance σ^2.

(b) the population standard deviation σ.

7. ***Vitamins*** ◆ To analyze the variation of vitamin supplement tablets, you randomly select and weigh 14 tablets. The results (in milligrams) are shown. Use a 90% level of confidence.

500.000	499.995	500.010	499.997	500.015	499.988	500.000
499.996	500.020	500.002	499.998	499.996	500.003	500.000

8. ***Cough Syrup*** ◆ You randomly select and measure the contents of 15 bottles of cough syrup. The results (in fluid ounces) are shown. Use a 90% level of confidence.

4.211	4.246	4.269	4.241	4.260	4.293	4.189	4.248
4.220	4.239	4.253	4.209	4.300	4.256	4.290	

9. ***Car Batteries*** ◆ The number of hours of reserve capacity of 18 randomly selected automotive batteries is shown. Use a 99% level of confidence. *(Adapted from Consumer Reports)*

1.70	1.60	1.94	1.58	1.74	1.60	1.86	1.72	1.38
1.46	1.64	1.49	1.55	1.70	1.75	0.88	1.77	2.07

Figure for Exercise 9 Figure for Exercise 10

10. ***Bolts*** ◆ You randomly select and measure 17 bolts. The results (in inches) are shown. Use a 95% level of confidence.

1.286	1.138	1.240	1.132	1.381	1.137	1.300	1.167	1.240
1.401	1.241	1.171	1.217	1.360	1.302	1.331	1.383	

Extending the Basics

Each exercise set ends with a group of exercises called *Extending the Basics*. These exercises go beyond the material presented in the section (they tend to be more challenging and are not required as prerequisites of subsequent sections).

Answers and Solutions

The answers to all odd-numbered exercises are given in the back of the book, and the worked-out solutions are available in the *Student's Solutions Manual.*

Section Exercise Sets

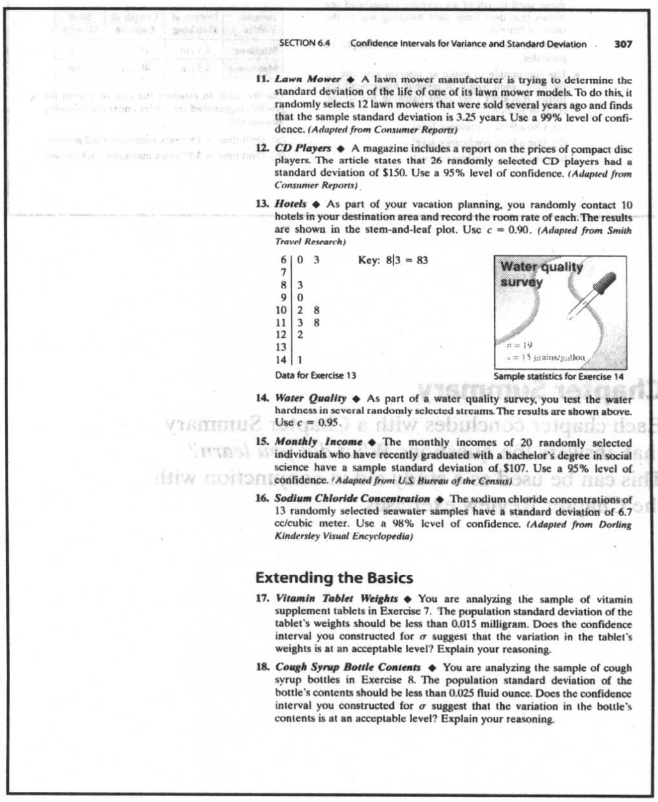
Each section concludes with a set of exercises carefully written to nurture student understanding and proficiency. They move from basic concepts and skill development to more challenging and interpretive problems.

Labeled Exercises

Most exercises are labeled for easy reference. For instance, exercises labeled Graphical Analysis ask students to use the graphs provided to answer the questions.

Paired Format

Almost all exercises are given in "paired format" so that the odd-numbered exercise, whose answer is given in the back of the book, is paired with an even-numbered exercise, whose answer is not given. This paired format is commonly used in mathematics texts, but is less common in statistics texts.



SECTION 6.4 Confidence Intervals for Variance and Standard Deviation 307

11. ***Lawn Mower*** ◆ A lawn mower manufacturer is trying to determine the standard deviation of the life of one of its lawn mower models. To do this, it randomly selects 12 lawn mowers that were sold several years ago and finds that the sample standard deviation is 3.25 years. Use a 99% level of confidence. *(Adapted from Consumer Reports)*

12. ***CD Players*** ◆ A magazine includes a report on the prices of compact disc players. The article states that 26 randomly selected CD players had a standard deviation of $150. Use a 95% level of confidence. *(Adapted from Consumer Reports)*

13. ***Hotels*** ◆ As part of your vacation planning, you randomly contact 10 hotels in your destination area and record the room rate of each. The results are shown in the stem-and-leaf plot. Use $c = 0.90$. *(Adapted from Smith Travel Research)*

6	0 3
7	
8	3
9	0
10	2 8
11	3 8
12	2
13	
14	1

Key: 8|3 = 83

Data for Exercise 13

Sample statistics for Exercise 14

14. ***Water Quality*** ◆ As part of a water quality survey, you test the water hardness in several randomly selected streams. The results are shown above. Use $c = 0.95$.

15. ***Monthly Income*** ◆ The monthly incomes of 20 randomly selected individuals who have recently graduated with a bachelor's degree in social science have a sample standard deviation of $107. Use a 95% level of confidence. *(Adapted from U.S. Bureau of the Census)*

16. ***Sodium Chloride Concentration*** ◆ The sodium chloride concentrations of 13 randomly selected seawater samples have a standard deviation of 6.7 cc/cubic meter. Use a 98% level of confidence. *(Adapted from Dorling Kindersley Visual Encyclopedia)*

Extending the Basics

17. ***Vitamin Tablet Weights*** ◆ You are analyzing the sample of vitamin supplement tablets in Exercise 7. The population standard deviation of the tablet's weights should be less than 0.015 milligram. Does the confidence interval you constructed for σ suggest that the variation in the tablet's weights is at an acceptable level? Explain your reasoning.

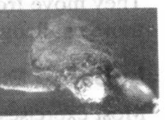
18. ***Cough Syrup Bottle Contents*** ◆ You are analyzing the sample of cough syrup bottles in Exercise 8. The population standard deviation of the bottle's contents should be less than 0.025 fluid ounce. Does the confidence interval you constructed for σ suggest that the variation in the bottle's contents is at an acceptable level? Explain your reasoning.

Chapter Case Study

Each chapter has a full-page case study featuring actual data from a real-world context and a series of thought-provoking questions that are designed to illustrate the important concepts of the chapter.

6 Case Study

WWW.NMFS.NOAA.GOV



Shell Lengths of Loggerhead Sea Turtles

The National Marine Fisheries Services is part of the National Oceanic and Atmospheric Administration. NMFS's programs support the conservation and management of living marine resources.

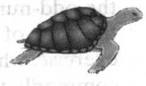
There are six species of sea turtles in the United States and all are protected as endangered or threatened species. Rarely does a hatchling sea turtle live to maturity. In fact, it is believed that only 1 in 10,000 hatchlings lives long enough to reproduce.

In a study by Hays and Marsh reported in the *Canadian Journal of Zoology* (75: 40–46, 1997), 71 loggerhead sea turtles were captured and measured off the coast of Britain. The shell lengths of the turtles are shown in the stem-and-leaf plot at the right.

Part of the purpose of the study was to estimate the growth rate of juvenile turtles. The turtles were hatched off the coast of Florida and their drifting time in the Atlantic Ocean was estimated to be between 1.8 and 3.75 years. From this and the fact that a typical hatchling has a shell length of 4.5 centimeters, Hays and Marsh estimated that juvenile loggerhead sea turtles grow at a rate of between 4.3 and 8.9 centimeters per year.

1	5 5 6 6 6 7 7 7 7 8 8 8 8 8 8 8 9 9 9 9 9
2	0 0 0 0 0 0 0 0 0 1 1 1 2 2 2 2 2 2 3 4 4
2	5 5 5 6 6 7
3	0 0 0 3 4
3	8
4	0
4	5 9
5	1 4
5	
6	0 1 1 4
6	
7	
7	5 8
8	
8	8 8
9	0 0
9	6
10	4

Key: 1|5 = 15 cm



Exercises

1. A loggerhead sea turtle is classified as a juvenile if its shell length is less than 40 centimeters. How many of the turtles in the sample were juveniles?
2. Use the sample to make a point estimate of the mean shell length of all juvenile loggerhead sea turtles that drift from their hatching site to the coast of Britain.
3. Find the standard deviation of the sample of juveniles.
4. Use the sample to make an interval estimate of the mean shell length of juvenile loggerhead sea turtles that drift from their hatching site to the coast of Britain.
   (a) Use a 90% confidence level.
   (b) Use a 95% confidence level.
   (c) Use a 99% confidence level.
5. How would your results have differed if you had used all the turtles in the sample instead of just the juvenile turtles? Explain your reasoning.
6. Complete the following table.

Juvenile Turtles	Length at Hatching	Length at Capture	Shell Growth
Minimum	4.5 cm	15 cm	?? cm
Maximum	4.5 cm	40 cm	?? cm

Use the table to estimate the rate of growth for juvenile loggerhead sea turtles under the following assumptions.

(a) Drift time = 1.8 years, minimum shell growth
(b) Drift time = 3.75 years, maximum shell growth

Chapter Summary

Each chapter concludes with a Chapter Summary that answers the questions *What did you learn?* This can be used as a study aid in conjunction with the Chapter Review exercises.

6 Chapter Summary

What did you learn?	*Review Exercises*
Section 6.1	
◆ How to find a point estimate and a maximum error of estimate	*1, 2*
◆ How to construct and interpret confidence intervals for the population mean	*3, 4*
◆ How to determine the required minimum sample size when estimating μ	*5, 6*
Section 6.2	
◆ How to interpret the t-distribution and use a t-distribution table	*7, 8*
◆ How to construct confidence intervals when $n < 30$ and σ is unknown	*9–14*
Section 6.3	
◆ How to find a point estimate for the population proportion	*15–20*
◆ How to construct a confidence interval for a population proportion	*21–26*
◆ How to determine the required minimum sample size when estimating a population proportion	*27, 28*
Section 6.4	
◆ How to interpret the chi-square distribution and use a chi-square distribution table	*29–32*
◆ How to use the chi-square distribution to construct a confidence interval for the variance and standard deviation	*33, 34*

Statistics: Uses and Abuses

Each chapter features an expanded discussion on how statistical techniques should be used, while cautioning students about common abuses. New exercises help students to apply their knowledge.

STATISTICS

Uses and Abuses

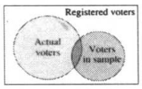
Uses

By now, you know that complete information about population parameters is often not available. The techniques of this chapter can be used to make interval estimates of these parameters so that you can make informed decisions.

From what you learned in this chapter, you know that point estimates of population parameters are rarely exact. Remembering this can help you make good decisions in your career and in everyday life. For instance, suppose the results of a survey tell you that 52% of the population plans to vote for a certain item on a ballot. You know that this is only a point estimate of the actual proportion that will vote for the item. If the interval estimate is $0.49 < p < 0.55$, then you know this means it is possible that the item will not receive a majority vote.

Abuses

Unrepresentative Samples There are many ways that surveys can result in incorrect predictions. When you read the results of a survey, remember to question the sample size, the sampling technique, and the questions asked. For example, suppose you want to know how a group of people will vote in an election. From the diagram at the right, you can see that even if your sample is large enough, it may not consist of actual voters.



Biased Survey Questions In surveys, it is also important to analyze the wording of the questions. For example, a question about rezoning might be presented as: "Knowing that rezoning will result in more businesses contributing to school taxes, would you support the rezoning ?"

Exercises

1. ***Unrepresentative Samples*** Find an example of a survey that is reported in a newspaper or magazine. Describe different ways that the sample could have been unrepresentative of the population.
2. ***Biased Survey Questions*** Find an example of a survey that is reported in a newspaper or magazine. Describe different ways that the survey questions could have been biased.

Real Statistics–Real Decisions

This new feature encourages students to think critically and make informed decisions about real-world data. Exercises guide students from interpretation to drawing conclusions.

Real Statistics–Real Decisions 315

PUTTING IT ALL TOGETHER

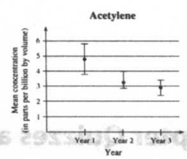
Real Statistics — Real Decisions

As part of the U.S. Environmental Protection Agency's (EPA) efforts to "protect human health and safeguard the natural environment," the EPA conducts the Urban Air Toxics Monitoring Program (UATMP). The program has gathered thousands of air samples and analyzed them for concentrations of over 50 different organic compounds, such as acetylene (used for cutting and welding metals, and in the production of neoprene rubber, plastics, and resins). The results from UATMP are used to gain insight into the effects of air pollution and if efforts to clean up the air are working.

EPA United States Environmental Protection Agency
www.epa.gov

For instance, using air samples from a major city, the EPA can analyze the results and estimate the mean concentration of acetylene in the air using a 95% confidence interval. They can then compare the interval to previous years' results to see if there are any trends and if there has been a significant change in the amount of acetylene in the air.

You work for the EPA and are asked to interpret the results shown in the graph at the right. The graph shows the point estimate for the population mean concentration and the 95% confidence interval for μ for acetylene over a three-year period. The data is based on air samples taken at one major city.



Exercises

1. ***Interpreting the Results***
   Consider the graph of the mean concentration levels of acetylene. For the following years, decide if there has been a change in the mean concentration level of acetylene. Explain your reasoning.
   (a) From Year 1 to Year 2
   (b) From Year 2 to Year 3
   (c) From Year 1 to Year 3
2. ***What Can You Conclude?***
   Using the results of Exercise 1, what can you conclude about the efforts to reduce the concentration of acetylene in the air?
3. ***How Do You Think They Did It?***
   How do you think the EPA constructed the 95% confidence interval for the population mean concentration of the organic compounds in the air? Do the following to answer the question (you do not need to make any calculations).
   (a) What sampling distribution do you think they used? Why?
   (b) Do you think they used the population standard deviation in calculating the maximum error of estimate? Why or why not? If not, what could they have used?